

PROGRESS 2

NLP

RQ1: Given a text, identify the inherent motivation as per Maslow hierarchy of needs.

Approach

- ▶ Finetune neural language models such as BERT, for **multi-label** classification task.
- ▶ 6 classes - {physiological, stability, love, esteem, spiritual growth, none}
- ▶ Dataset :

Link for the paper: <https://arxiv.org/pdf/1805.06533.pdf>

Link for the Video: <https://vimeo.com/285805584>

Link for the data: <https://uwnlp.github.io/storycommonsense/>

Defining advanced, basic and none needs.

Basic Needs : Sentences which are mapped to Maslow categories: [physiological ,love,stability] by annotators.

Advanced Needs: Sentences which are mapped to Maslow categories: [esteem,spiritual growth] by annotators.

None: Sentences which are mapped to Maslow category none by annotators.

Datasets

Original Dataset

sentence	
maslow	
esteem	5187
love	4502
none	2543
physiological	2277
spiritual growth	2360
stability	3724

Train Set

sentence	
maslow	
esteem	4149
love	3602
none	2034
physiological	1822
spiritual growth	1888
stability	2979

Test Set

sentence	
maslow	
esteem	1038
love	900
none	509
physiological	455
spiritual growth	472
stability	745

sentence	
maslow	
advanced_needs	6037
basic_needs	8403
none	2034

Data Level 1 Preprocessing

- ✓ Step 1: Figured out Number of samples in each category of subreddit
- ✓ Step 2: Extracted only samples of Subreddit Depression
- ✓ Step 3: Merged Title and Text as Title_Text(Lets call it sentence)

```
Index(['unique_id', 'Title', 'Text', 'Subreddit', 'title_text'], dtype='object')
```

	Text
Subreddit	
Anxiety	86243
BPD	38216
autism	7142
bipolar	41493
depression	258496
mentalhealth	39373
schizophrenia	17506

✓ Step 4: Removing samples which have sentence with total number of words as 1, 2, 3,.....,14

1 Word Sentence: empty \n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n, talk

2 Word Sentence: Testing Hello, AAAAAAAAAAAAAA AAAAAAAAAAAAAA, meme <https://imgflip.com/i/28uv8g>

3 Word Sentence: Help me . , Inner voices [caption] (<http://imgur.com/WLbIIO...>), Ha ha Hahahhaha

4 Word Sentence: Day 4... I'm done..., tmrw ...is the day. , Depression is... [Depression is...](<https://imgur.com/gallery/i4po9>)

Eg: Some of the samples have images which tells about the user being depressed

<https://imgur.com/gallery/i4po9>, <https://imgur.com/n6Pb6H2>, <https://i.imgur.com/c8DLrxX.jpg>, <https://imgur.com/gallery/bezPJ>

5 Word Sentence: i am a mess I, So I turned 18... Yay, WHAT A LOVELY DAY a, Bad day today. Enough said.

6 Word Sentence: I slept 18 hours today. Goddammit., Procrastination + Perfectionism = NOT GOOD, Henlo H e n l o

7 Word Sentence: Stephen Hawking died Ree,

Similarly we had noise as above examples and few of them doesn't showcase any need according to our understanding. So we assumed to ignore these sentences and remove sentences which are less than 15 Word Length.

✓ Step 5: Non English and Duplicates posts are removed. URLs were also removed.

The final Reddit Depression Dataset had 242,434 Sentences

As future work URL's and Images can be analysed.

* 20 to 200 words sentence post to be taken based on graph

Classification or Labeling the depression posts with maslow needs

- Clarification – BERT model finetuned on SocialChem Dataset labelled with maslow needs. The dataset is made from commonsense short stories (Mostafazadeh et al., 2016) and hand annotated. The dataset has a total of ? 14440 texts labelled as a unique Maslow Need. Go to Slide 5 – for more details.
- Objective – Depressions posts are unlabelled with maslow needs. Manually impossible/tedious to tag them. Prior works have used finetuned models for annotations, so that's why we are also doing it.
- Perform prediction.
- Quality Check – 50 randomly picked posts (25 + 25) to understand the type of samples marked as False positive or False Negative.
- Next Steps – tweaking threshold for prediction ?

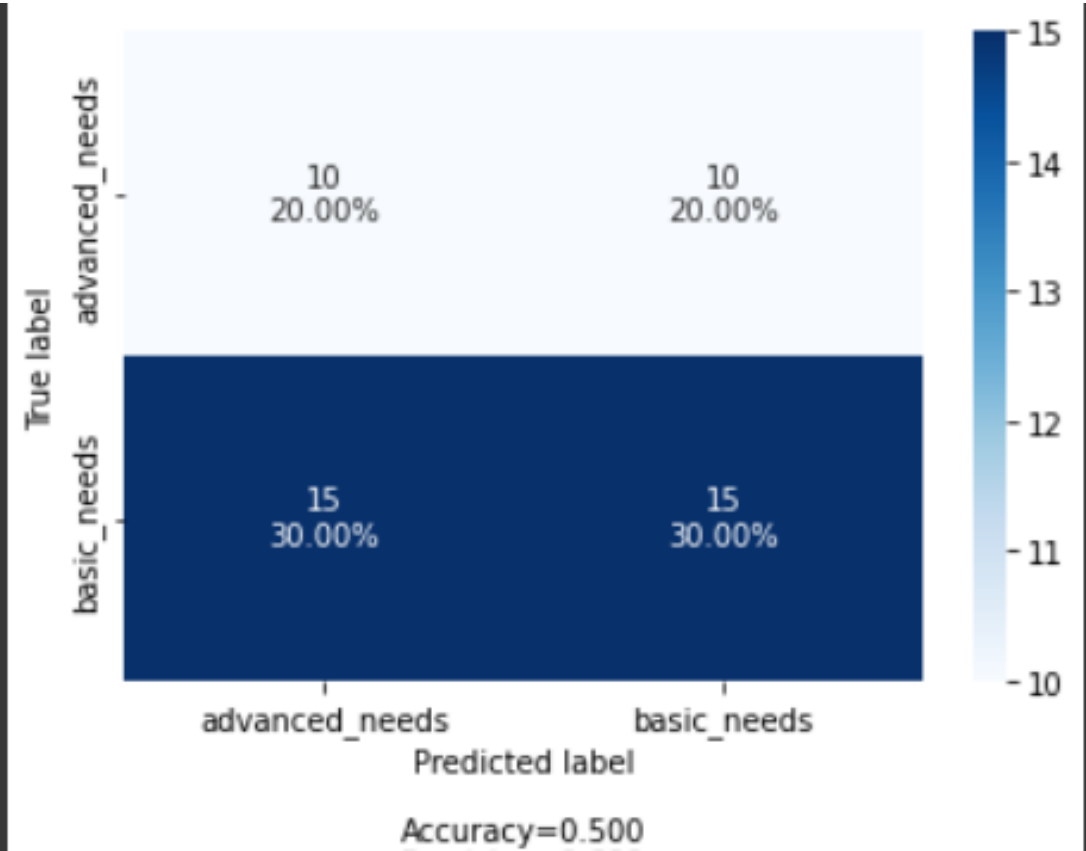
Quality Checking after Model Annotations

final_prediction	our_label_majority	
advanced_needs	advanced_needs	10
	basic_needs	15
basic_needs	advanced_needs	10
	basic_needs	15

AFTER ANNOTATIONS

```
print(classification_report(df['our_label_majority'], df[
```

	precision	recall	f1-score	support
advanced_needs	0.40	0.50	0.44	20
basic_needs	0.60	0.50	0.55	30
accuracy			0.50	50
macro avg	0.50	0.50	0.49	50
weighted avg	0.52	0.50	0.51	50



QUALITY CHECK ANALYSIS

Classification or Labelling of M1 (Basic and Advanced Needs) by M_basic and M_advanced

- M_basic was trained on 8403 basic Needs Text (physiology, love, stability) from SocialChem Dataset to label the Reddit basic need sentences identified by M1 as [physiology, love, stability].
- M_advanced was trained on 6037 advanced Needs Text (spiritual growth, esteem) from SocialChem Dataset to label the Reddit advanced need sentences identified by M1 as [spiritual growth, esteem].
- Now from 50 sampled posts taken for annotation we filtered out sentences upon which the model M1 and we agreed upon for classifying it further to 5 Maslow needs by M_basic and M_advanced. Later these were again manually annoated by us.

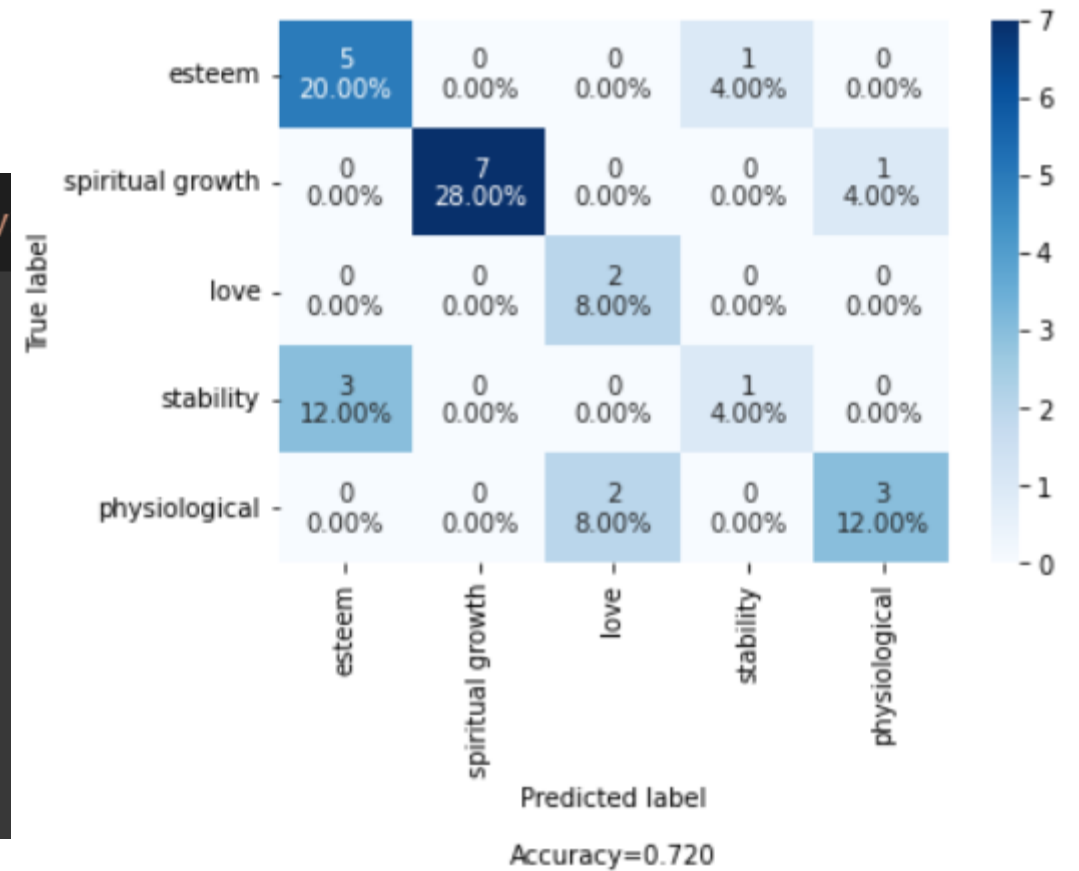
Quality Checking after Model Annotations

final_prediction_maslow	our_label_majority2	
esteem	esteem	5
	spiritual growth	3
love	love	7
physiological	physiological	2
	stability	2
spiritual growth	esteem	1
	spiritual growth	1
stability	love	1
	stability	3

AFTER ANNOTATIONS

```
[54] print(classification_report(correct_match['our_label_majority
```

	precision	recall	f1-score	support
esteem	0.62	0.83	0.71	6
love	1.00	0.88	0.93	8
physiological	0.50	1.00	0.67	2
spiritual growth	0.50	0.25	0.33	4
stability	0.75	0.60	0.67	5
accuracy			0.72	25
macro avg	0.68	0.71	0.66	25
weighted avg	0.74	0.72	0.71	25



QUALITY CHECK ANALYSIS

Few Reaction Posts

- waking up depressed again... it's not even been 30 minutes yet and i am already crying.
- stop the world, I want to get off. i am no longer enjoying this ride I have not for quite sometime. I think I would like to get off now.
- everything has been fine for three months until i lost a bolt while i was working on my bike. i just stood there staring at the ground thinking about how i should kill myself later today.
- waking up depressed again... it's not even been 30 minutes yet and i am already crying.

Suggestions

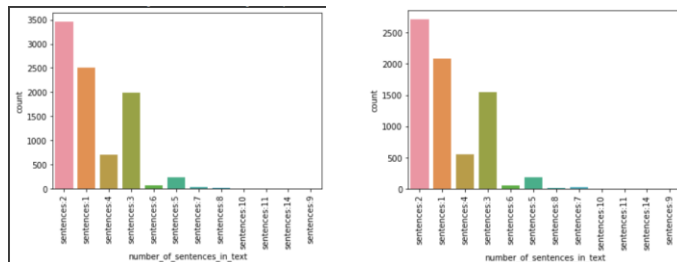
- Fix word length 20-200 – check for 50 random samples – if no reaction posts, then move to the next prediction step. If there are still posts talking about depression/reaction, try filter with "depres".
 - This will filter reaction posts.
- Check Sentence Distribution for the dataset after above filtering step. **If high number of sentences, then sequence modeling can be performed.**
- Perform prediction. Do a quality check – basic vs advanced labeling (refer Slide 16)
- Confusion Matrix?

Suggestions

- Look at the number of sampled for each bucket:
- Bucket Size:
 - Create a box plot Title_text_words for 250K after Level 1 Preprocessing which tells us quantiles and outliers. (4 intervals <word length>)
 - Pick atleast 20-30 from each bucket and check manually to see if a certain bucket is useless <if it includes reaction post>
- Any 2 people work on 300 samples of Reddit.
 - Perform Inter annotation Agreement

Level 2 Preprocessing

- The depression sentences after level 1 preprocessing were 242434
- On removing sentences which doesn't have word length in range [20,200] the total number of depression sentences were 152748.
- And on applying filter to remove reaction sentences using the keyword "depres" the total number of depression sentences were 99843.
- WHAT WE DID: We did a Level 2 Preprocessing on these 15K randomly sampled sentences
- If Level 2 Filter was applied on 15K randomly sampled sentences then:
- 15000 -> 8984 -> 7172

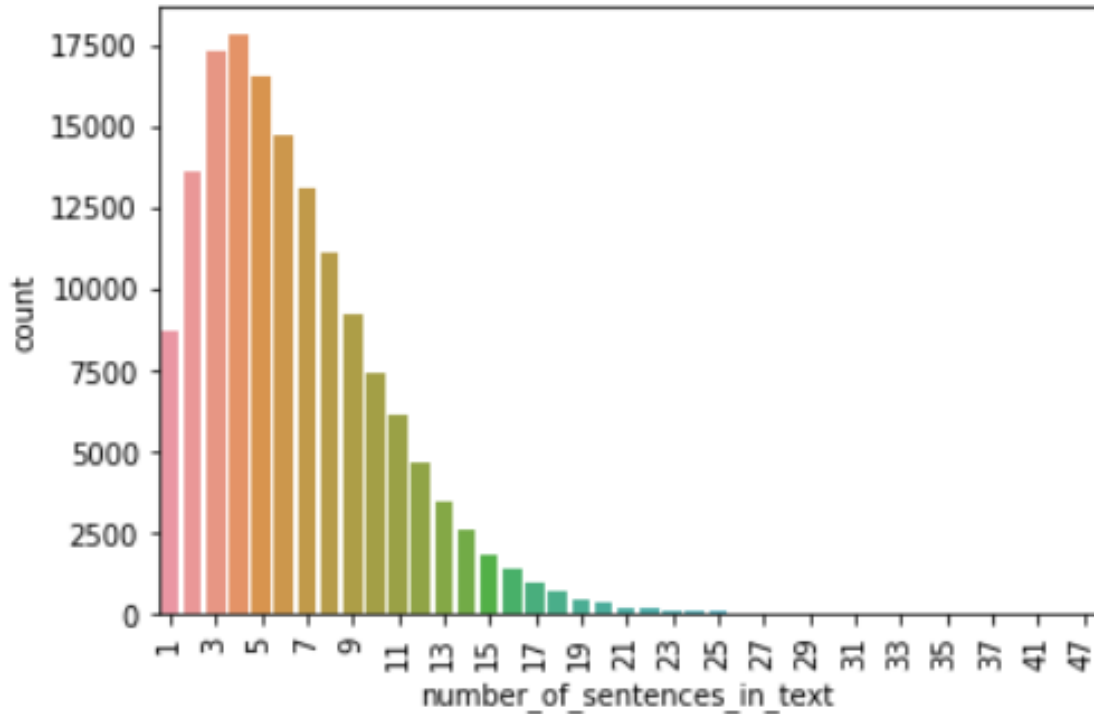


Annotation Scheme

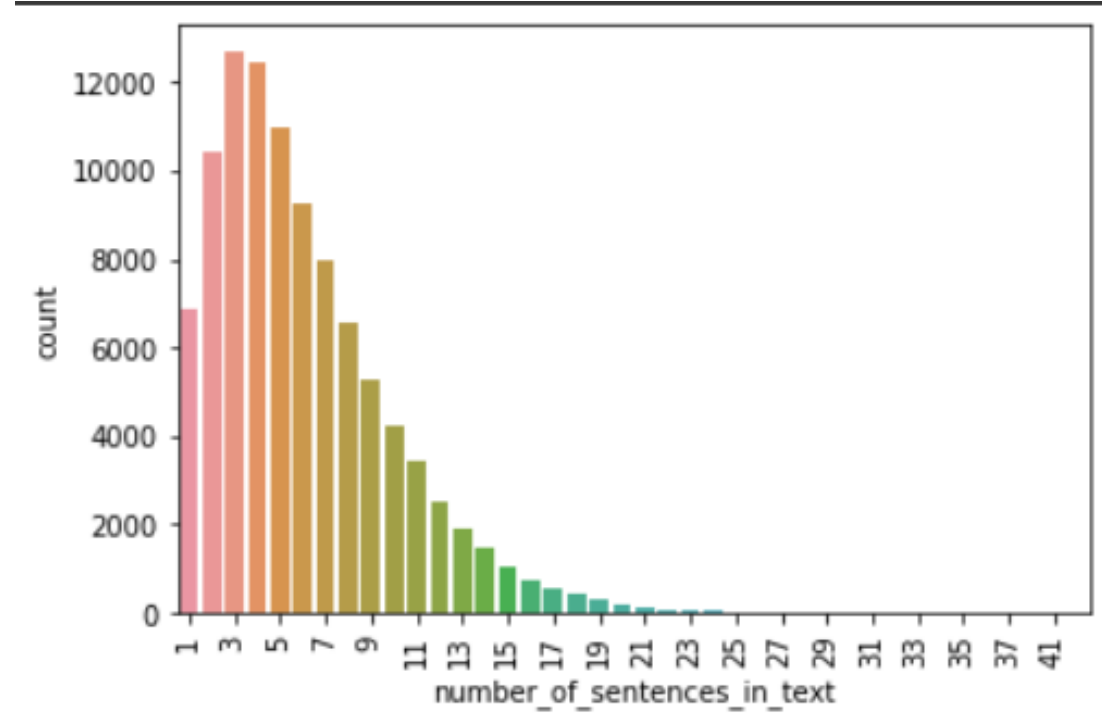
- Identify the entities/agents of conversation involved.
- Identify the overall context of conversation – is it about a family gathering, is it your job related/work related, it is school?
- Interpret the statement w.r.t entities and overall context.
- If the need is still not clear/ambiguous, label it ambiguous.

Overall Preprocessed Depression dataset after Level 2 Preprocessing

Number of Sentences in a Title_Text



Before Applying Filter to remove reaction words

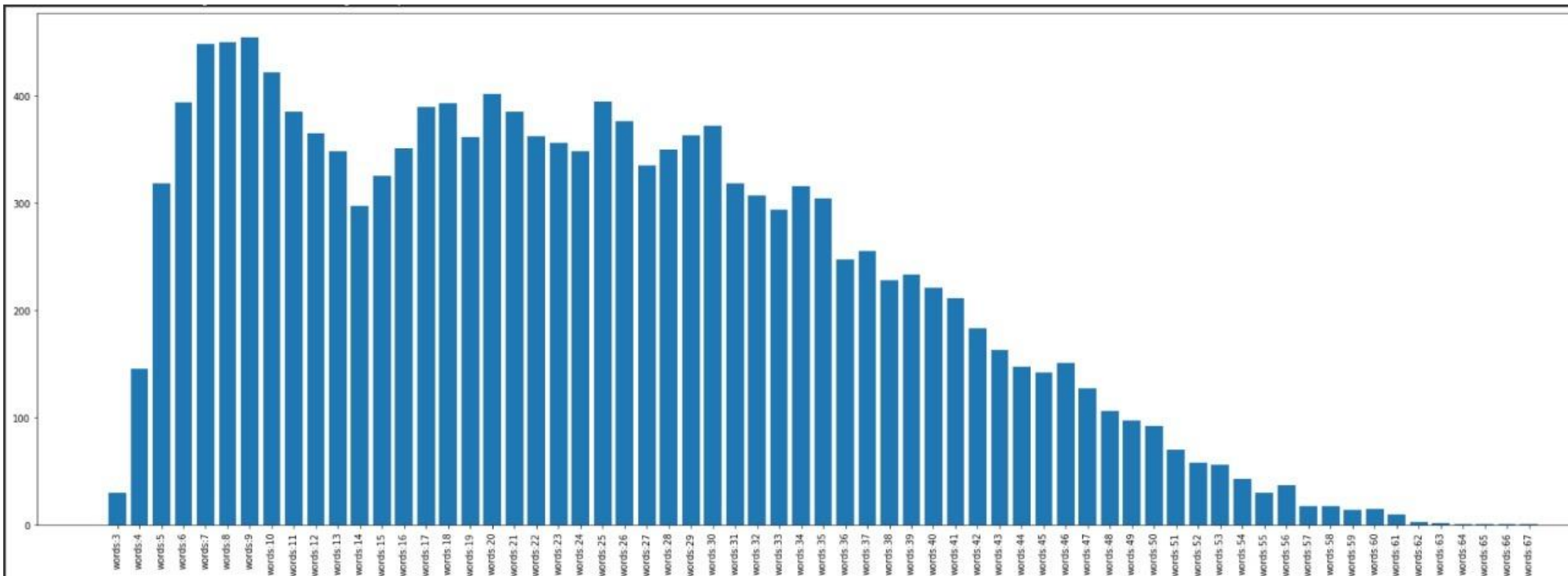


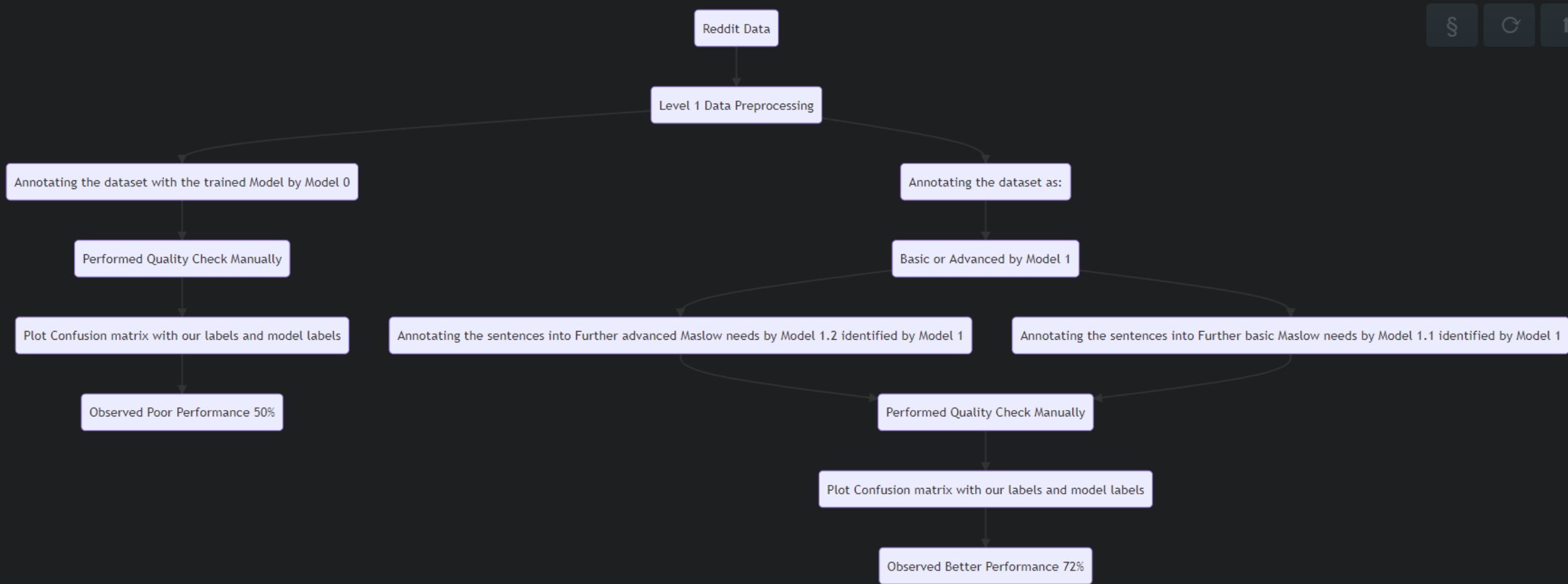
After Applying Filter to remove reaction words

Data Processing:

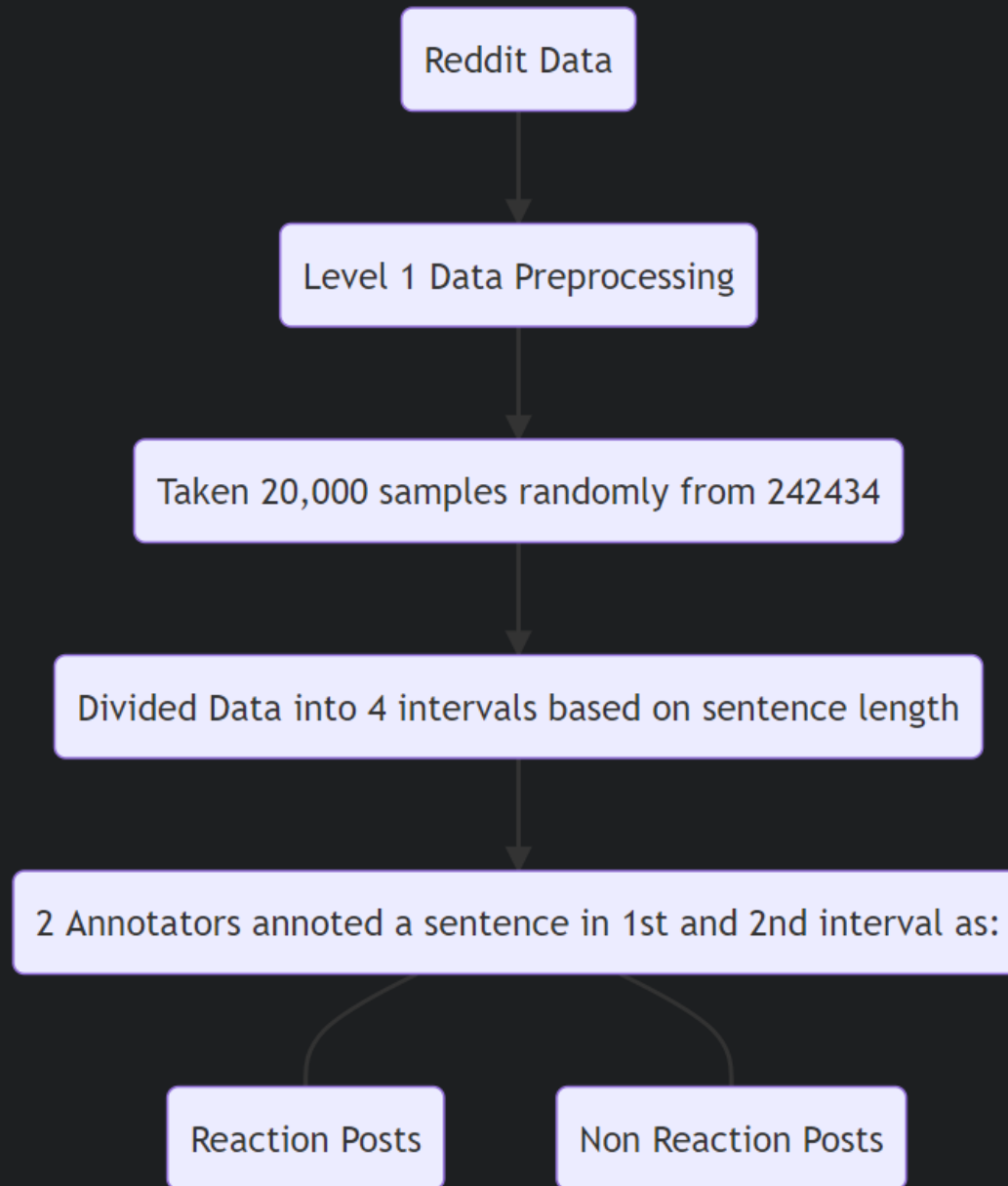
- Depression samples taken from subreddit r/depression. No manual labeling of posts. Assumed that since it belonged to depression thread, it is posted by a depressed person.
- Depression samples (258322). Due to computational constraints, pulled a subset of ~15K. This 15K was randomly picked after first level of filtering
- Filtering – Every post should have title + text. Non English and Duplicates posts are removed. Posts with <3 words in text are also removed. URLs were also removed. This was done on 258322 set. After filtering, we are left with 228247.
- Insert word distribution plot for filtered 228K set. -- so that we can logically select the subset of 3K. [Title, Text, Title + text]
- Fix a threshold for #words after looking at plots, then randomly select 15K.
- Do a Quality check on 50 posts.

Word Distribution in SocialChem Dataset





Identified Reaction Posts which didnt indicate any Maslow Need in the Reddit Dataset. So:



How did you do decide posts as Reaction or Non- Reaction

- Reason of Marking as Reaction Post: Noticed that reaction posts are not indicating Maslow Need. Hence, they need to be removed.
- When a sentence indicates some sort of mental or physical reaction towards his/her condition then it is marked as Reaction Posts.
- When a sentence indicates a need of a person then we mark it the Maslow need which it indicates instead of marking it as Reaction or Non Reaction. This was done as there was a question which had not been answered above yet.

Word Distribution Intervals for the sampled Reddit Dataset

	unique_id	number_of_words_in_title_Text
count	20000.000000	20000.000000
mean	135137.129600	209.977950
std	6175.688736	244.876012
min	124460.000000	4.000000
25%	129774.750000	69.000000
50%	135131.500000	136.000000
75%	140492.250000	258.000000
max	145814.000000	4476.000000

Dividing into Intervals based on Word Length

- first interval: [4-69] word length sentences.
- second interval: [70-136] word length sentences.
- third interval: [137-258] word length sentences.
- fourth interval: [259-4476] word length sentences.

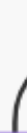
M1 Model Trained on Social Chem dataset =>basic/advanced<=



92 labelled sentences given by Ajinkya of interval 2



Split it into:



Train Set =>72<=

Test Set =>20<=

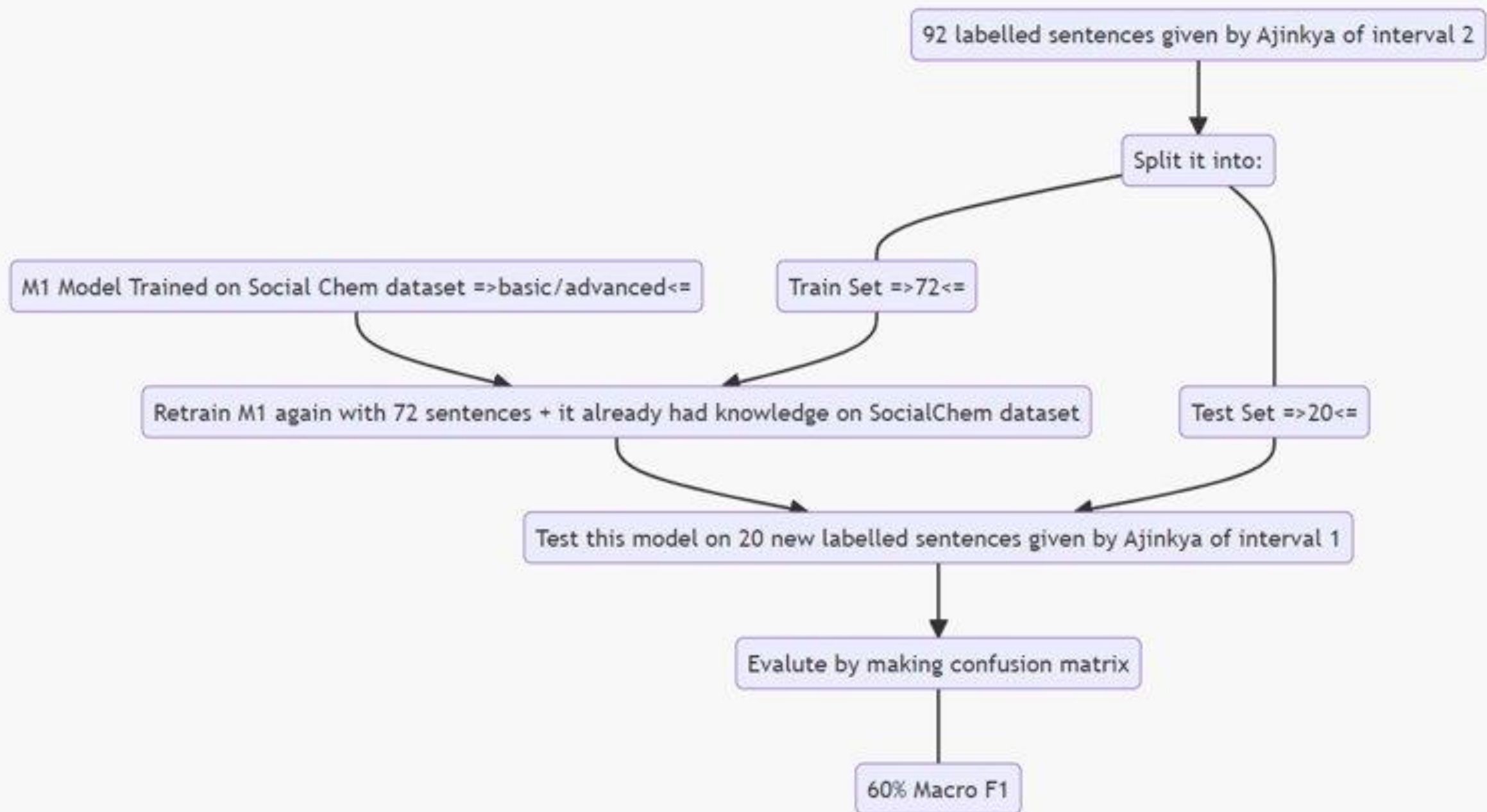


Test M1 using train set & making confusion matrix



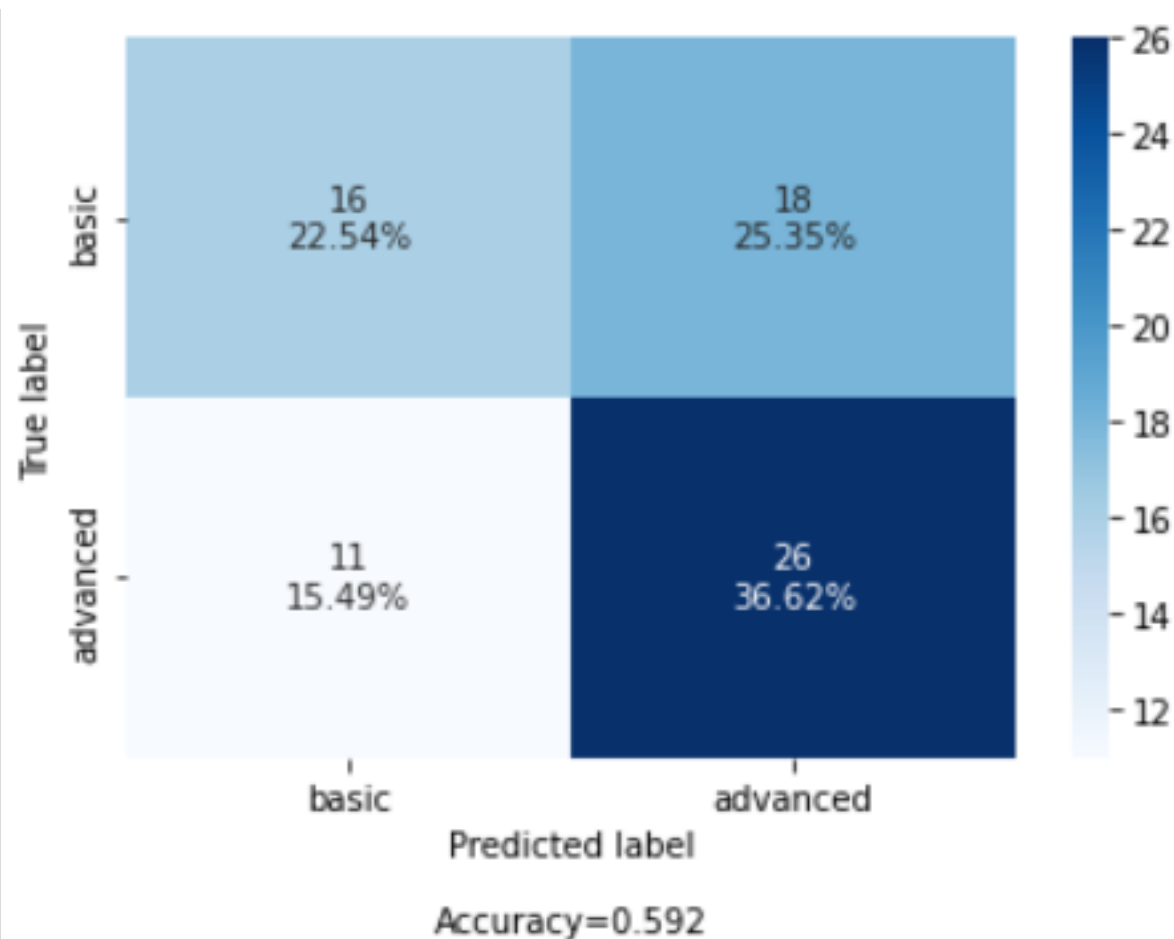
59% Macro F1

Simulation of Few Shot Learning:



Testing Model M1 on Ajinkya's Annotated posts(72)

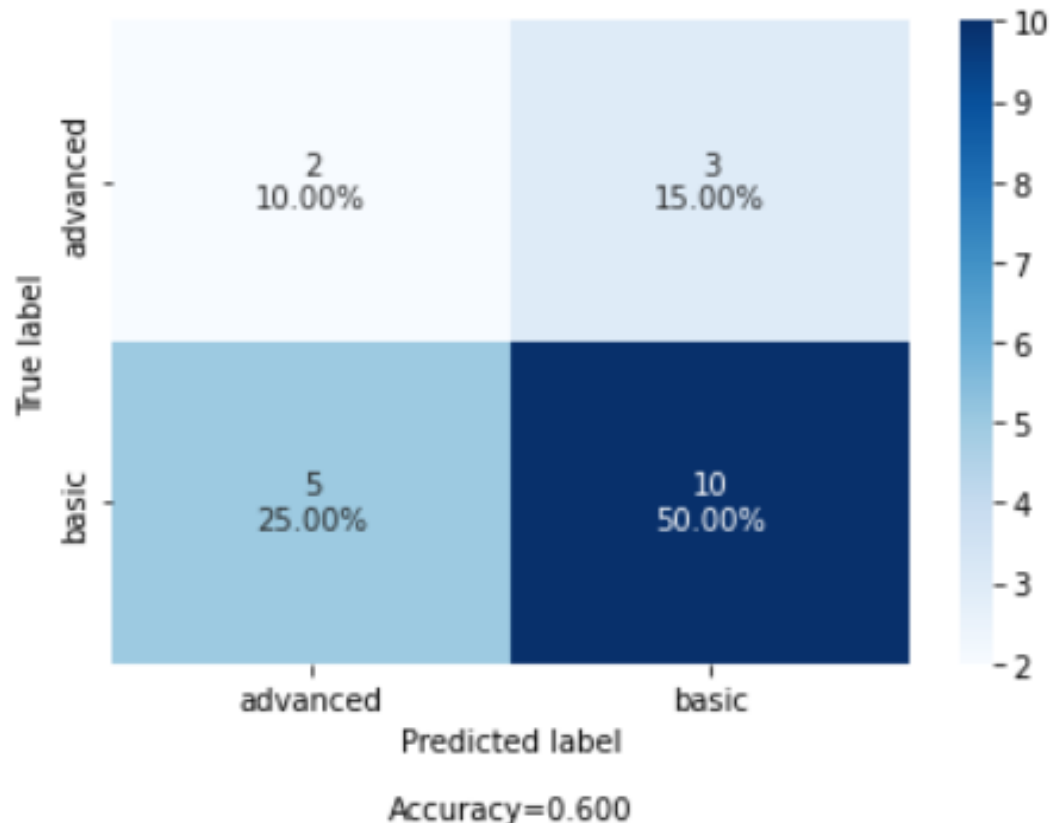
- M1 labelled 72 posts marked by Ajinkya as basic or advanced.



	precision	recall	f1-score	support
advanced_needs	0.59	0.47	0.52	34
basic_needs	0.59	0.70	0.64	37
accuracy			0.59	71
macro avg	0.59	0.59	0.58	71
weighted avg	0.59	0.59	0.59	71

Simulation of Few Shot Learning

- Retrained M1 with Ajinkya Labelled 72 posts and tested this model on 20 test set posts which was also annotated by Ajinkya.



	precision	recall	f1-score	support
advanced_needs	0.29	0.40	0.33	5
basic_needs	0.77	0.67	0.71	15
accuracy			0.60	20
macro avg	0.53	0.53	0.52	20
weighted avg	0.65	0.60	0.62	20

Further Goals

- Lets convolute the process and try to use GPT-3. Reason: Easier Approach and more novelty. Read the paper <https://arxiv.org/pdf/2210.08604.pdf>
- Develop a prompt for GPT-3 which would provide good results.
- Decide the Post which should be fed into GPT-3.

Prompt 1

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. List the needs which are mentioned in the below post.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- Needs mentioned in the post: -Safety -Belongingness -Esteem -Self-actualization

Prompt 2

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain the needs which are mentioned in the below post.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- The post expresses the need for belongingness, safety, and self-actualization. Belongingness is the need to feel connected to others and to be accepted and loved by them. Safety is the need to feel secure and protected from harm. Self-actualization is the need to reach one's full potential in life.

Prompt 3

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain which of these five needs are mentioned in the below post.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- The needs mentioned in this post are belongingness, esteem, and self-actualization. The person is expressing a need for belongingness by feeling envious of others in relationships and feeling left out. They are expressing a need for esteem by feeling the need to be accepted by others and questioning why they don't deserve love and acceptance. Finally, they are expressing a need for self-actualization by wanting to be happy and feeling like their current situation is preventing them from achieving that.

Prompt 4

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. List the needs which are mentioned in a post. Return none if there is no identifiable need.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- Belongingness, esteem, safety

Prompt 5

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain the needs which are mentioned in the below post. Return none if there is no identifiable need.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- Belongingness: Help with rejection, depression, suicide issues Esteem: I feel so depressed, like why me? why don't I deserve this? Self-Actualization: I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel.

Prompt 6

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain which of these five needs are mentioned in the below post. Return none if there is no identifiable need.
- Post: Help with rejection, depression, suicide issues I am 22 and have never had a girlfriend and have been rejected countless times. I see couples all around me and I feel so depressed, like why me? why don't I deserve this? I feel like killing myself. I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel. I don't know what to do?
- Belongingness, esteem, and safety.

6 Prompts which were framed

- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. List the needs which are mentioned in the below post.
- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain the needs which are mentioned in the below post.
- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain which of these five needs are mentioned in the below post.
- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. List the needs which are mentioned in a post. Return none if there is no identifiable need.
- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain the needs which are mentioned in the below post. Return none if there is no identifiable need.
- Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain which of these five needs are mentioned in the below post. Return none if there is no identifiable need.

Comparision Between 6 Prompts for a Single Post

- Physiological, safety, belongingness, esteem.
- The post expresses the need for belongingness, safety, and self-actualization. Belongingness is the need to feel connected to others and to be accepted and loved by them. Safety is the need to feel secure and protected from harm. Self-actualization is the need to reach one's full potential in life.
- The needs mentioned in this post are belongingness, esteem, and self-actualization. The person is expressing a need for belongingness by feeling envious of others in relationships and feeling left out. They are expressing a need for esteem by feeling the need to be accepted by others and questioning why they don't deserve love and acceptance. Finally, they are expressing a need for self-actualization by wanting to be happy and feeling like their current situation is preventing them from achieving that.
- Belongingness, esteem, safety
- Belongingness: Help with rejection, depression, suicide issues Esteem: I feel so depressed, like why me? why don't I deserve this? Self-Actualization: I hate it when people say you don't need a relationship to make you happy because the people saying that have had relationships so they don't know how I feel.
- Belongingness, esteem, and safety.

Data Worked with to Test the Prompts

- Number of Samples of Reddit Data which was manually annotated by Ajinkya and Vivek: 80
- Distribution of words(in general): [70,136]
- Distribution of unique words: [52,96]

Few Questions

- 124472 : Prompt 2 gave definitions of all categories as well as reason
- 124485: Doubt in reasoning Prompt 3 gave
- 124508: Check both reasonings
- 124511: Both Prompts gave valid explanations but one in descriptive fashion and other in a brief manner
- Prompt 1 gave None for some sentences even we can find a need getting displayed. Eg: 124516
- 124517: Felt like the user didn't want any party

- After comparing of 6 prompts we decided to get 60 responses(20 from each team member) on 3 prompts namely:
 1. Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. List the needs which are mentioned in a post. Return none if there is no identifiable need.
 2. Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain the needs which are mentioned in the below post.
 3. Task: Human needs are of five types physiological, safety, belongingness, esteem and self-actualization. Explain which of these five needs are mentioned in the below post.

Insights

- After checking the predictions of GPT₃ on the 80 sentences we identified that predictions was correct but it didn't establish the causality i.e is Depression causing a need (lack of belongingness) or lack of belongingness causing depression.
- If self-harm is mentioned in the post the machine returned safety need. But the person is responsible for lack of safety i.e he is talking about self-harm. Therefore here lack of safety isn't causing depression.
- Therefore depression could be completely unrelated of person's surrounding or situation.
- In case of GPT₃: The prediction's appears to be correct but may not be until we understand and answer the question "why this output?". "Is there a trend for all the things which were labelled with certain label?". If answered these then we can identify the shortcomings of GPT₃ i.e what is it capturing and what it is not capturing.
- Even though the need seems valid the task is to understand the interpretation.

Tasks:

- Find no. of cases where there was no correlation between with the need and depression.
- Make a classification task:
 - For classifying "No. Of posts which are expressing self harm vs safety"
 - Then remove those posts which contain self-harm by not considering them as having safety need.
- Try to find underlaying pattern across all 80 needs. "Why was it annotated belongingness?"
 - Firstly take the definitions given by model for all the needs.
 - Then take outputs corresponding to Prompt 2 and Prompt 3.
 - Pick all the posts marked with a need (eg: belongingness) and check the underlaying pattern.
 - If found confusion between self-actualization and esteem categories, then mention that we found them to be interchangeable and mention them as advanced needs.
 - If 5 annotators are doing it then do mention the agreement scores.
- Look at questions in PHQ-9 and try to make prompts more accurate (prompt engineering) such that we can establish causality. Later try the engineered prompts on the 80 samples to check GPT₃'s predictions.
- Find any manual for checking depression symptoms from psychologists and check with the doctor if we are following the right approach for solving the problem.
- Perform Self Verification
- Propose a new metric for self verification

- After modifications:

Insights:

1. After checking the predictions of GPT₃ on the 80 sentences we identified that predictions was correct but it didn't establish the causality I.e is Depression causing a need (lack of belongingness) or lack of belongingness causing depression.
2. If self-harm is mentioned in the post the machine returned safety need. But the person is responsible for lack of safety I.e he is talking about self-harm. Therefore here lack of safety isn't causing depression.
3. Therefore depression could be completely unrelated of person's surrounding or situation.
4. In case of GPT₃: The prediction's appears to be correct but may not be until we understand and answer the question "why this output?". "Is there a trend for all the things which were labelled with certain label?". If answered these then we can identify the shortcomings of GPT₃ I.e what is it capturing and what it is not capturing.
5. Even though the need seems valid the task is to understand the interpretation. If there is a unfulfilled need or belongingness.

Tasks:

1. Find no. of cases where there was no correlation between with the need and depression.
2. Make a classification task:
 - For classifying "No. Of posts which are expressing self harm vs safety"
3. Try to find underlying pattern across all 80 needs. "Why was it annotated belongingness?"
 - Firstly take the definitions given by model for all the needs.
 - Then take outputs corresponding to Prompt 2 and Prompt 3.
 - Pick all the posts marked with a need (eg: belongingness) and check the underlying pattern.
 - If found confusion between self-actualization and esteem categories, then mention that we found them to be interchangeable and mention them as advanced needs.
 - If 5 annotators are doing it then do mention the agreement scores.
4. Look at questions in PHQ-9 and try to make prompts more accurate (prompt engineering) such that we can establish causality. Later try the engineered prompts on the 80 samples to check GPT3's predictions.
5. Find any manual for checking depression symptoms from psychologists and check with the doctor if we are following the right approach for solving the problem.
6. Perform Self Verification
7. Propose a new metric for self verification



THANK YOU
